**WHAT IS CLAIMED IS:**

1. A speech synthesis system, which synthesizes speech using time series data of formant parameters (including a formant frequency and a formant bandwidth) estimated based on a speech production model, the speech synthesis system comprising determining the correspondence of formant parameters between adjacent frames using dynamic programming.

2. The speech synthesis system of Claim 1, wherein in determining the correspondence of the formant parameters, a connection cost $d_c(F(n), F(n+1))$ and a disconnection cost $d_d(F(k))$ are obtained using the equations:

$$d_c(F(n), F(n+1)) = \alpha\left|F_f(n) - F_f(n+1)\right| + \beta\left|F_i(n) - F_i(n+1)\right|$$

$$\begin{aligned} d_d(F(k)) &= \alpha\left|F_f(k) - F_f(k)\right| + \beta\left|F_i(k) - \varepsilon\right| \\ &= \beta\left|F_i(k) - \varepsilon\right| \end{aligned}$$

where $\alpha$ and $\beta$ are predetermined weight coefficients, $F_f(n)$ is a formant frequency in the $n^{th}$ frame, that $F_i(n)$ is a formant intensity in the $n^{th}$ frame and $\varepsilon$ is a predetermined value, and the resultant $d_c(F(n), F(n+1))$ and $d_d(F(k))$ are used as costs for grid point shifting in dynamic programming.

3. The speech synthesis system of Claim 2, wherein for two adjacent frames in which exists a formant which has no counterpart to be connected,

a formant having the same frequency as that of the disconnected

formant in one of the frames and an intensity of 0 is located in the other frame and

the two adjacent frames are connected by interpolation of frequencies and intensities of both the formants according to a smooth function.

4. The speech synthesis system of Claim 2, wherein the formant intensity $F_i(n)$ is calculated using

$$F_i(n) = \begin{cases} 20\log_{10}\left(\dfrac{1+e^{-\pi F_b(n)/F_s}}{1-e^{-\pi F_b(n)/F_s}}\right) & \text{, if formant} \\ 20\log_{10}\left(\dfrac{1-e^{-\pi F_b(n)/F_s}}{1+e^{-\pi F_b(n)/F_s}}\right) & \text{, if anti-formant} \end{cases}$$

where $F_b(n)$ is a formant bandwidth in the $n^{th}$ frame and $F_s$ is a sampling frequency.

5. The speech synthesis system of Claim 3, wherein a vocal tract transfer function including a plurality of formants is implemented by a cascade connection of a plurality of filters and

wherein when a formant which has no counterpart to be connected exists in the adjacent frames and thus the connection of the filters needs to be changed,

a coefficient and an internally stored data of the filter in question are copied into another filter and

the first filter is then over written with a coefficient and an internally stored data of still another filter or initialized to predetermined values.

6. The speech synthesis system of Claim 4, wherein a vocal tract transfer function including a plurality of formants is implemented by a cascade connection of a plurality of filters and

wherein when a formant which has no counterpart to be connected exists in the adjacent frames and thus the connection of the filters needs to be changed,

a coefficient and an internally stored data of the filter in question are copied into another filter and

the first filter is then over written with a coefficient and an internally stored data of still another filter or initialized to predetermined values.

7. A speech analysis method, in which a sound source parameter and a vocal tract parameter of a speech signal waveform are estimated by using a glottal source model including an RK voicing source model, the speech analysis method comprising the steps of:

extracting an estimated voicing source waveform using a filter which is constituted by the inverse characteristic of an estimated vocal tract transfer function;

estimating a peak position corresponding to a GCI (glottal closure instance) of the estimated voicing source waveform with higher accuracy at closer time intervals than that with the sampling period by applying a quadratic function;

synthesizing the GCI with a sampling position in the vicinity of the estimated peak position and thereby generating a voicing

source model waveform; and

time-shifting the generated voicing source model waveform with higher accuracy at closer time intervals than that with the sampling period by means of all pass filters and thereby matching the GCI with the estimated peak position.

8. A speech analysis method, in which a voicing source parameter and a vocal tract parameter of a speech signal waveform are estimated by using a glottal voicing source model such as an RK model or a model defined as a modified model thereof, the speech analysis method comprising the steps of:

extracting an estimated voicing source waveform using filters which are constituted by the inverse characteristic of an estimated vocal tract transfer function; and

assuming the first harmonic level as H1 and the second harmonic level as H2 in DFT (discrete Fourier transformation) of the estimated voicing source waveform and estimating an OQ (open quotient) from a value for HD defined as HD=H2-H1.

9. The speech analysis method of Claim 8, wherein for estimating the OQ, the relation:

$$OQ=3.65HD-0.273HD^2+0.0224HD^3+50.7$$

is used.